Integrative analysis of spontaneous CLL regression highlights genetic and microenvironmental interdependency in CLL

Kwok, Marwan Cheng Kuang; Oldreive, Ceri; Rawstron, Andy C; Goel, Anshita; Papatzikas, Grigorios; Jones, Rhiannon; Drennan, Samantha; Agathanggelou, Angelo; Sharma-Oates, Archana; Evans, Paul; Smith, Edward; Dalal, Surita; Mao, Jingwen; Hollows, Robert; Gordon, Naheema; Hamada, Mayumi; Davies, Nick; Parry, Helen; Beggs, Andrew; Munir, Talha

DOI:
10.1182/blood.2019001262

Integrative analysis of spontaneous CLL regression highlights genetic and microenvironmental interdependency in CLL

Running Title: Natural history and mechanism of spontaneous CLL regression

Marwan Kwok,[1,2,3,*] Ceri Oldreive,[1,*] Andy C. Rawstron,[3] Anshita Goel,[1,4] Grigorios Papatzikas,[1,4] Rhiannon E. Jones,[5] Samantha Drennan,[6] Angelo Agathanggelou,[1] Archana Sharma-Oates,[1,4] Paul Evans,[3] Edward Smith,[1] Surita Dalal,[3] Jingwen Mao,[1] Robert Hollows[1], Naheema Gordon,[1] Mayumi Hamada,[1] Nicholas Davies,[1] Helen Parry,[2,7] Andrew D. Beggs,[1] Talha Munir,[3] Paul Moreton,[8] Shankara Paneesha,[9] Guy Pratt,[1,2] A. Malcolm R. Taylor,[1] Francesco Forconi,[6] Duncan M. Baird,[5] Jean-Baptiste Cazier,[1,4] Paul Moss,[2,7,**] Peter Hillmen,[3,10,**] and Tatjana Stankovic[1,**]

[1] Institute of Cancer and Genomic Sciences, University of Birmingham, UK
[2] Centre for Clinical Haematology, Queen Elizabeth Hospital Birmingham, Birmingham, UK
[3] Haematological Malignancy Diagnostic Service, St. James's University Hospital, Leeds, UK
[4] Centre for Computational Biology, University of Birmingham, UK
[5] Division of Cancer and Genetics, School of Medicine, Cardiff University, Cardiff, UK
[6] Cancer Sciences Unit, University of Southampton, Southampton, UK
[7] Institute of Immunology and Immunotherapy, University of Birmingham, UK
[8] Department of Haematology, Pinderfields General Hospital, Wakefield, UK
[9] Department of Haematology, Birmingham Heartlands Hospital, Birmingham, UK
[10] Section of Experimental Haematology, University of Leeds, Leeds, UK

* M.K. and C.O. contributed equally to this work
** P.M., P.H. and T.S. contributed equally to this work

Corresponding authors:

Marwan Kwok, MB ChB, MRCP (UK), PhD
Institute of Cancer and Genomic Sciences, University of Birmingham
Institute for Biomedical Research, Vincent Drive, Edgbaston, Birmingham, UK
m.kwok@bham.ac.uk

Tatjana Stankovic, MD, PhD
Institute of Cancer and Genomic Sciences, University of Birmingham
Institute for Biomedical Research, Vincent Drive, Edgbaston, Birmingham, UK
t.stankovic@bham.ac.uk

Article Type:	Regular Article
Scientific Section:	Lymphoid Neoplasia, Clinical Trials and Observations
Text word count:	4000
Abstract word count:	249
No. of figures:	7 (plus 8 supplemental)
No. of tables:	5 (plus 6 supplemental)
No. of references:	54 (plus 6 supplemental)
Keywords:	Spontaneous remission, tumor heterogeneity, BCR signaling, whole exome sequencing, RNA-seq

This manuscript contains a data supplement.

Key Points

1. Spontaneously regressed tumors are composed of a formerly proliferating CLL clone that has transitioned into a quiescent state.

2. A change in microenvironmental stimulation on an indolent genomic background state underpins clonal attrition in spontaneous CLL regression.

Abstract

Spontaneous regression is a recognized phenomenon in chronic lymphocytic leukemia (CLL) but its biological basis remains unknown. We undertook a detailed investigation of the biological and clinical features of 20 spontaneous CLL regression cases incorporating phenotypic, functional, transcriptomic and genomic studies at sequential timepoints. All spontaneously regressed tumors were *IGHV*-mutated with no restricted *IGHV* usage or B-cell receptor (BCR) stereotypy. They exhibited shortened telomeres similar to non-regressing CLL indicating prior proliferation. They also displayed low Ki-67, CD49d, sIgM expression and IgM signaling response but high CXCR4 expression indicating low proliferative activity associated with poor migration to proliferation centers, with these features becoming increasingly marked during regression. Spontaneously regressed CLL displayed a transcriptome profile characterized by downregulation of metabolic processes as well as *MYC* and its downstream targets compared to non-regressing CLL. Moreover, spontaneous regression was associated with reversal of T-cell exhaustion features including reduced PD-1 expression and increased T-cell proliferation. Interestingly, archetypal CLL genomic aberrations including *HIST1H1B* and *TP53* mutations and del(13q14) were found in some spontaneously regressing tumors, but genetic composition remained stable during regression. Conversely, a single case of CLL relapse following spontaneous regression was associated with increased BCR signaling, CLL proliferation and clonal evolution. These observations indicate that spontaneously regressing CLL appear to undergo a period of proliferation before entering a more quiescent state, and that a complex interaction between genomic alterations and microenvironment determines disease course. Together, the findings provide novel insight into the biological processes underpinning spontaneous CLL regression, with implications for CLL treatment.

Introduction

Spontaneous tumor regression, manifested by a sustained reduction or disappearance of the malignant clone without treatment, is an uncommon but intriguing phenomenon.[1-6] In CLL, spontaneous disease regression is estimated to occur in 1-2% of patients,[7,8] although CLL cells do remain at a low level despite the resolution of clinical features.[7-9] Several case reports of spontaneous CLL regression have been published,[9-11] as well as 2 case series which together describe a total of 19 cases.[7,8]

Spontaneous tumor regression has been linked in some cases to infection, blood transfusion or secondary malignancy, and it has been suggested that these may reflect induction of a pro-inflammatory state.[2,11] However, these associations are uncommon and as such the etiology of spontaneous regression is unclear. A study of 9 cases of spontaneously regressing CLL showed that tumors lacked CD38 and ZAP-70 expression and all 7 cases that underwent *IGHV* sequencing revealed an *IGHV*-mutated genotype with *IGHV3* usage in 6 of these.[8]

Improved understanding of the biological basis underlying spontaneous CLL regression is important for elucidating CLL pathobiology and could help to inform therapeutic management. Multiple lines of evidence now suggest that the clinical outcomes for patients with CLL is shaped both by the biological properties of the malignant clone and its interaction with the microenvironment.[12] Disease progression is often accompanied by increased CLL migration to proliferation centers with resultant BCR engagement and clonal proliferation.[13-15] Moreover, the emergence of mutations and genomic complexity may result in the outgrowth of CLL cells with greater proliferative and survival advantage.[16-18] These changes, facilitated by impaired tumor immunity and a microenvironment conducive to CLL proliferation and apoptotic evasion,[19-23] may shift the balance of CLL turnover in favor of clonal expansion. As such we postulated that spontaneous CLL regression may reflect the integration of a range of processes including BCR signaling and proliferation, apoptosis, genetic heterogeneity and host immunity.

To evaluate this hypothesis, we analyzed sequential samples from 20 patients who had undergone spontaneous CLL regression. Our analysis revealed unique features providing insight into the biology of this phenomenon.

Patients, Material and Methods

Patient samples

We reviewed all 1425 patients with untreated CLL who attended 4 United Kingdom hemato-oncology centers between 2010 and 2016. We identified complete spontaneous CLL regression on the basis of a sustained reduction in absolute lymphocyte count (ALC) to below 4 x 10^9/L with complete resolution of CLL-related symptoms, anemia, thrombocytopenia and clinically detectable adenopathy. Identification of partial spontaneous regression was based on sustained reduction of lymphocytosis by ≥50% from the highest level, with regressing nodal disease. We excluded subjects with concurrent infections or second malignancies, or who were receiving myelosuppressive or immunosuppressive therapies, including systemic corticosteroids, for any indication immediately preceding or coinciding with the onset of CLL regression. In two individuals the onset of CLL regression predated the diagnosis of a second malignancy by ≤5 years (CLL18–19), which could impact on immune responses contributing to CLL regression. However, upon subsequent analysis no distinctive features were apparent between these two and other patients, thus, they were incorporated into the rest of the cohort.

For comparison, CLL-treatment naïve, Binet stage-, age- and sex-matched indolent CLL patients; age- and sex-matched progressive CLL patients and healthy controls were recruited. Indolent cases were further CMV-status matched for T-cell studies. To reflect real-world CLL populations, groups were not matched for *IGHV* mutational status. However, subgroup analysis excluding the *IGHV* unmutated indolent cases did not affect the results. Indolent disease and healthy controls were recruited locally, while progressive samples were sourced from multicenter trials and obtained immediately before treatment (supplemental Table 1). Indolent CLL was defined as Binet stage A disease with a ≥2 year lymphocyte

doubling time monitored over ≥5 years. Samples were obtained with written patient consent and institutional ethical approval.

Cell separation

CLL cells (>95% CD19+ CD5+ cells; supplemental Table 2), for all non-FACS based assays, were isolated using a MACS procedure (Miltenyi Biotec, Surrey, UK), as detailed in supplemental methods.

Flow cytometry

Extracellular and intracellular staining procedures are detailed in supplemental methods and supplemental Table 3.

Nucleic acid extraction

Telomere length, genomic and transcriptomic studies were performed on DNA and mRNA extracted from isolated CLL cells.

Telomere length, telomerase and replicative senescence assays

XpYp chromosome-specific single telomere length analysis (STELA) was performed as previously described.[24] Telomerase activity was assessed using the Telo*TAGGG* telomerase PCR enzyme-linked immunosorbent assay (Roche, Sussex, UK). Replicative senescence was evaluated using β-galactosidase staining (Cell Signaling Technology, Beverly, MA). Details are in supplemental methods.

***IGH* sequencing**

IGH sequencing was performed as described in supplemental methods.[25]

SNP array analysis, whole exome sequencing and RNA-seq

Single-nucleotide polymorphism (SNP) array analysis and whole exome sequencing (WES) was performed on CLL cells, with isolated granulocyte[26] or T-cell fraction as germline control. There was no detectable *IGHV* rearrangement in the germline fractions (supplemental Figure 1). Details of sample preparation, bioinformatics analysis and result validation are provided in supplemental methods. Raw WES and RNA-seq data is available in the NCBI SRA (Bioproject ID: PRJNA535508; URL: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA535508).

Results

Spontaneous disease regression occurs in <2% of patients with CLL and is seen exclusively in tumors with mutated *IGHV*

We identified 20 individuals who fulfilled the criteria for spontaneous CLL regression, a prevalence of 1.4%, which approximates well with reported frequencies.[7,8] Eleven subjects underwent complete spontaneous regression (CLL01–11) and 6 partial regression (CLL12–17). In 2 individuals, onset of CLL regression predated the diagnosis of a second malignancy by ≤5 years (CLL18–19). In addition, one patient (CLL20) exhibited complete CLL regression for 14 years but subsequently relapsed (Figure 1; Table 1).

Patients were followed for a median of 15.2 years (range 3.6–30). Among complete regressors, CLL regression occurred over a median of 8.6 years (range 5.3–26). At the time of study, lymphocytosis and CLL-related clinical features had been absent for a median of 2.7 years (range 0.9–12) and consistent with the significantly reduced tumor load, T-cell and NK numbers resembled normal levels (Table 2). There was no association between onset of regression and initiation of non-CLL related medication for any patient (Table 1). None of the patients were taking non-prescriptive medication known to influence CLL course, immediately preceding or coinciding with the onset of CLL regression, nor were they frequent green tea consumers. However, the potential contribution of other lifestyle factors, holistic or dietary, to regression cannot be discounted. All subjects except one (CLL07) were alive at the time of analysis. Importantly, a residual B-lymphocyte population with a CLL phenotype was

identifiable by flow cytometry in all the regression samples accounting for a median of 92.5% of total B-cells (range 71.6–99.8%; Figure 2A; Table 2).[27]

The sequence of the clonal *IGHV* rearrangement was determined in 19 tumors and found to be *IGHV*-mutated (M-CLL) in all instances (Table 2). A range of *IGHV* usage was observed, including *IGHV3* in 12 cases, *IGHV4* in 5 cases, and a single case each of *IGHV1* and *IGHV2*. None exhibited stereotyped BCR usage.

Spontaneous regressed CLL tumors have undergone substantial cell division but gradually acquired a low proliferative phenotype

We analyzed the phenotypic features of the residual CLL clone in spontaneously regressed (n=17), indolent (n=54) and progressive (n=40) tumors. We also examined for phenotypic changes occurring during spontaneous regression by comparing paired diagnostic and regression samples (n=5). The relative accumulation or regression of a tumor reflects the balance between cellular proliferation and apoptosis.[13,28] We therefore focused initially on the proliferative activity of the CLL clone. Spontaneously regressed CLL contained significantly fewer Ki-67-expressing tumor cells (<1%) than either indolent (median 2.5%; $P \leq 0.05$) or progressive (median 5%, $P \leq 0.0001$) CLL, indicating a very low level of clonal proliferation (Figure 2B). CLL cells that have recently exited proliferation centers and stopped receiving proliferation stimuli have low CXCR4 levels, which rapidly recover in the circulation.[29-31] Consistent with 3.46-fold lower Ki-67 expression, we found significantly higher (1.12-fold; $P \leq 0.05$) CXCR4 positivity on the CLL cells in spontaneously regressed compared to indolent cases (Figure 2C, supplemental Figure 2). In contrast, no significant differences were observed in the surface expression of CXCR5, CCR6 and CCR7 (supplemental Figure 3).

To investigate the proliferation history of spontaneously regressed CLL, we compared Ki-67 expression from paired diagnosis and regression samples. Ki-67 expression significantly reduced from diagnostic levels in all 5 patients with concomitant increased surface CXCR4 in 3 cases, suggesting the possibility of a lower proliferation rate during regression (Figure 2D). Additionally, the mean telomere length of spontaneously regressed CLL was found to be

comparable to that of non-regressing indolent CLL and was generally <5 kb (Figure 2E), suggesting previous proliferation consistent with prior diagnosis of bona-fide CLL.[24,32] The clonal composition of the tumor was also assessed by *IGH* sequencing and revealed a single clonal population in each case (Table 2). This remained stable between diagnosis and regression, as evidenced by identical *IGHV-D-J* sequence and similar telomere length distribution (supplemental Figure 4). Taken together, our results demonstrate that spontaneously regressed CLL tumors comprise clones that have previously undergone substantial cell division but have subsequently acquired a phenotype of low proliferation.

CLL tumors undergoing spontaneous regression display weak BCR signaling responses but retain high levels of Bcl-2

Clonal attrition is likely to occur when the rate of CLL cell death exceeds that of CLL proliferation. Given the dominant role of positive BCR signaling in driving CLL proliferation, we asked whether CLL regression might reflect an attenuation of BCR signaling responses. All spontaneously regressed tumors were ZAP-70 negative (supplemental Figure 5). We analyzed BCR levels and signaling responses in 14 sIgM/sIgD+ spontaneously regressed tumors (Figure 3). Anti-IgM mediated responses were lower in indolent M-CLL than in *IGHV*-unmutated (UM)-CLL, consistent with previous studies.[33-36] However, responses were further reduced to undetectable levels in spontaneously regressed CLL (Figure 3B). In contrast, substantial IgD responses were retained, although they appeared lower in spontaneously regressed CLL compared to indolent M-CLL and UM-CLL (Figure 3C). Indeed, IgD signaling accounted for almost all residual signaling upon combined IgM/IgD stimulation in spontaneously regressed CLL (Figure 3D), with reduced IgM signaling being associated with low sIgM expression (Figure 3E-F). All regressed tumors expressed similar sIgM levels to M-indolent CLL which were lower than UM-indolent CLL (Figure 3E-F). Class switch recombination, evident by IgG expression was responsible for the negligible sIgM expression only in a single case (CLL19; data not shown). Thus, the sIgM levels potentially suggest an anergic state for the majority of regression cases. Furthermore, comparison of sequential

spontaneous regression samples revealed 5.01-fold increased CXCR4 and significant ($P \leq 0.005$) 4.61-fold reduced surface immunoglobulin expression (Figure 2D & 3G), suggesting reduced antigen engagement within proliferation centers and positive BCR signaling over time, consistent with development of a more quiescent state, potentially anergy.[35,37] Corroborating these findings, spontaneously regressed CLL exhibited other anergic features such as 2-fold higher levels of basal Erk phosphorylation (pERK) than healthy B-cells (Figure 4A).[38] Also basal pErk[38] and B-cell inhibitory receptor LAIR1[39] levels were significantly elevated in regressed CLL compared to indolent M-CLL and UM-CLL ($P \leq 0.05$), whilst basal Akt phosphorylation was similar (Figure 4A-B).

CLL proliferation is dependent on the migration of CLL cells from the peripheral blood to the lymph nodes. We evaluated CD49d,[40,41] CD62L,[42] CD38,[41] and ROR1[43]expression, molecules important for migration. While no significant difference was observed in ROR1, CD38 or CD62L levels, surface CD49d expression was significantly lower between spontaneously regressed CLL and indolent tumors, being expressed on <2% of tumor cells in most cases (Figure 4C-F, supplemental Figure 2). Concurrent with the attenuation of BCR signaling during regression, regressed samples exhibited reduced CD38 (2.44-fold) and CD49d (6.93-fold) expression compared to diagnosis but only in cases with diagnostic expression on >5% CLL cells (Figure 4G). As CD49d was the only downregulated chemokine receptor, reduced CD49d-mediated trafficking may contribute to CLL regression.

We next assessed the expression of apoptosis-related proteins in CLL cells. Relative to indolent CLL, spontaneously regressing CLL displayed significantly higher FasR expression at both mRNA (logFC 0.98, $p<0.001$) and protein level (Figure 4H, supplemental Figure 2), implicating a potential role of the extrinsic apoptotic pathway in spontaneous regression. In contrast, comparable levels of the anti-apoptotic proteins Bcl-2 and Mcl-1 was seen in spontaneously regressed and indolent tumors (Figure 4I-J). Spontaneously regressed CLL continue to express Bcl-2 (Figure 4I), which could potentially contribute to the persistence of residual disease along with the concurrent upregulation of chemokines such as IL-8/CXCL8 and CXCL2 (Table 3).[44-46]

We hypothesized that the very low rate of proliferation in spontaneously regressed CLL clones could also be due to replicative senescence secondary to telomere erosion. To investigate this, we measured telomerase activity and beta-galactosidase expression, a marker of replicative senescence, within tumor cells. However, telomerase activity was higher in spontaneously regressed CLL tumors relative to non-regressing tumors and no evidence of beta-galactosidase staining was observed (supplemental Figure 6). As such there is no evidence to support replicative senescence as an explanation for decreased CLL proliferation during spontaneous regression.

Taken together, these results indicate that reduced BCR signaling, coupled with potentially curtailed migration to proliferation centers leads to a very low proliferation rate in spontaneously regressed tumors. Consistent with a low metabolic, quiescent state, RNA-seq analysis of 15 spontaneously regressed CLL tumors demonstrated downregulation of pathways involved in RNA metabolism, mRNA translation, protein biosynthesis and oxidative phosphorylation compared to age-matched indolent M-CLL (n=17) (Figure 5A-B; supplemental Table 4). Furthermore, *MYC*; an important gene promoting survival and proliferation[33] was one of the top 5 down-regulated genes in regressors and was associated with reduced Myc target gene expression (Figure 5C). Upon both principal component analysis (PCA) and multi-dimensional unsupervised hierarchical clustering analysis of the entire RNA-seq dataset which also included age-matched progressive M-CLL, UM-CLL and healthy controls, spontaneously regressed CLL segregated into a cluster that overlapped with the indolent M-CLL cluster (Figure 5D-E). This is consistent with spontaneously regressed CLL having a transcriptomic profile that bears the closest resemblance to indolent M-CLL. Together, this data potentially indicates that that prior to regression, an indolent genomic state is maintained which may render susceptibility to regression following an alteration of microenvironmental stimulation.

Spontaneous CLL regression occurs in the context of low genomic complexity and intraclonal equilibrium

The genomic landscape of individual CLL clones is believed to be a major determinant of biological behavior.[16,17] Analyses of spontaneously regressed CLL tumors at various timepoints (n=27) identified 12-57 somatic SNV per sample (supplemental Table 5). Furthermore, at the point of maximal spontaneous regression, validated variants of potential functional significance (n=18) revealed a pattern of low mutational burden and an absence of genomic complexity among spontaneously regressed CLL, with a median of 1 somatic copy number variation (CNV, range 0-3) and 2 single nucleotide variations (SNV, range 0-5) per case (Figure 6A; Table 4). The most frequent CNV, del(13q14), was present in 11 tumors (61%), and was clonal in most cases, consistent with it being an early genomic event and a likely driver of the initial proliferative phase before regression onset. Mutations involving known CLL drivers were seen in 4 cases (22%): a subclonal *HIST1H1B* mutation in one case, and clonal or subclonal *TP53* mutations in 3 cases (Figure 6A). All 3 *TP53* alterations were reported missense variants, considered to be damaging by both Polyphen and SIFT in silico prediction tools. Variant allelic frequencies (VAFs) ranged from 11.83-100%. The alteration TP53_g13804G>A has been reported multiple times in CLL, while all other alterations have been reported in other malignancies (COSMIC: http://cancer.sanger.ac.uk/cosmic). Only one (CLL13) occurred concurrently with 17p deletion (supplemental Table 6).

To determine the pattern of subclonal dynamics during spontaneous CLL regression, we analyzed genomic data from sequential diagnostic and regression samples (n=5), including 2 complete regression cases (CLL05, CLL06). Intraclonal equilibrium was a predominant feature at the point of regression in 4 subjects, indicated by similar subclonal representation across sequential timepoints and no major clonal selection or evolution (Figure 6B-C). A degree of clonal evolution involving 3 genes was evident in a single case (Figure 6D). These observations indicate that in similarity to MBL,[47] clonal expansion of spontaneously regressed CLL is absent, even in cases harboring driver mutations. This is consistent with the notion that mutational burden alone is insufficient to drive malignancy and that sustained clonal expansion requires an interplay between mutational burden, potential genomic instability and microenvironmental signals. To investigate this, we revisited our RNA-seq dataset. In addition

to downregulation of metabolic and proliferative processes, several genes involved in DNA repair and telomere maintenance were upregulated in spontaneously regressing CLL compared to M-indolent CLL (Table 5). The upregulation of the DNA repair genes was confirmed by qPCR (supplemental Figure 7). Notwithstanding the presence of genomic aberrations, these features may potentially contribute to the absence of genome instability in spontaneously regressing CLL.

Tumor immunity represents another mechanism that could suppress CLL proliferation in spontaneously regressing disease. CLL regression was accompanied by significantly reduced PD-1, LAG3 and CD160 expression on CD8+ T-cells compared to indolent cases ($P \leq 0.05$), whilst CD244 levels were unchanged (Figure 7A, supplemental Figure 2). Furthermore, PD-1 levels decreased from diagnosis on both CD4+ (2-fold) and CD8+ (3-fold) T-cells (Figure 7B). T-cell proliferation also significantly ($P \leq 0.05$) increased in comparison to their respective diagnostic timepoint and indolent CLL (Figure 7C). This indicates reversal of T-cell exhaustion possibly due to reduced tumor burden, which may result in improved cytotoxicity against CLL cells.[21] Furthermore, established CLL is associated with impaired immune function including elevated numbers of T- and NK-cells,[19-22] and regression was associated with normalization of T-cell and NK-cell number (Table 2, supplemental Figure 8). However, T-cell subset distribution was not significantly altered in spontaneously regressed CLL compared to indolent CLL, and CD4:CD8 inversion persisted in some cases.

Enhanced co-stimulation can overcome inhibition of BCR signaling *in vitro* and disease acceleration may be driven by clonal evolution

To determine the clinical significance of the residual disease that remains after spontaneous CLL regression, we examined whether the quiescent state of BCR unresponsiveness in spontaneously regressed CLL could be reversed. We compared BCR responses with and without CD40L and IL-4 co-culture as this has previously been shown to reverse clonal anergy.[48,49] Provision of co-stimulation with CD40L/IL-4 resulted in a recovery of BCR responses (elevation of pErk 13.76-fold, pAkt 9.38-fold and pSyk 14.31-fold) in

spontaneously regressed CLL (Figure 7D), indicating that BCR unresponsiveness is reversible, and likely due to chronic antigenic stimulation in the absence of tissue IL-4. This substantiates the proposition that CLL cells in regression cases were confined to the peripheral circulation for prolonged periods. We compared sequential CLL20 samples, a case characterized by CLL relapse following complete spontaneous regression. Disease relapse coincided with recovery of BCR responses and CLL proliferation (Figure 7E-F). Moreover, relapse was accompanied by clonal evolution in which the subclone harboring *TP53*, *KLRC1* and *RPGR* mutations, dominant at the time of spontaneous regression, was replaced by another with *ATM*, *MDC1* and *PRP21* mutations at relapse (Figure 6C). Collectively, these findings demonstrate that CLL cells from cases of spontaneous CLL regression remain susceptible to enhanced co-stimulation and this may potentially represent one mechanism of disease progression following clonal evolution.

Discussion

In this study, we present the largest cohort of spontaneous CLL regression cases analyzed to date. To investigate the natural history and biological processes underpinning spontaneous CLL regression, these cases underwent extensive characterization. This involved sequential timepoints, and integration of flow cytometric, functional, transcriptomic and genomic analyses, which distinguishes this study from previously published reports.

Our results support a model in which spontaneously regressed CLL arises from a CLL clone that has undergone extensive prior proliferation, but subsequently transitioned to a limited proliferative state with hyporesponsiveness to IgM BCR stimulation. The residual CLL clone is likely to be concentrated within peripheral blood due to reduced expression of CD49d which mediates adhesion and transendothelial migration, important for CLL cell trafficking to proliferation centers. However, it ultimately remains viable because of sustained Bcl-2 expression. Clonal quiescence, possibly together with increased DNA repair activity, enables maintenance of genome stability and intraclonal equilibrium despite the presence, in some

cases, of biologically relevant mutations. Such a model highlights the interdependence of genetic and microenvironmental factors in determining the clinical course of patients with CLL (Figure 7G). We postulate that under conditions of low genomic complexity and genome stability, mutations alone, in the absence of BCR signaling and other microenvironmental interactions, are insufficient for CLL progression or to sustain disease. Thus, it appears that an indolent genomic state is a required background in most CLLs, on which an alteration of microenvironmental/BCR stimulation may lead to regression and conversely, subsequent relapse in combination with genetic alterations. Both intrinsic and extrinsic factors are likely to contribute to these changes and therefore probably determine the balance between CLL quiescence and proliferation. The order of events and the causative factors may be case-dependent and could depend on unknown environmental factors.

This model explains the efficacy of BCR signaling inhibitors in producing durable remissions in the frontline setting, where genomic complexity is typically low even in the presence of genetic alterations such as *TP53* mutations.[17,18] In these cases, the effect of treatment with BCR signaling inhibitors bears resemblance to spontaneous CLL regression. In contrast, in relapsed or refractory CLL with high intrinsic levels of genomic complexity and instability, BCR signaling inhibition alone may be insufficient in preventing disease relapse arising from clonal evolution, and therapeutic strategies targeting both genetic and microenvironmental drivers may be required. Our findings also underscore the significance of detectable residual CLL following clinical regression, the outcome of which is dependent upon the dynamic interplay between factors that promote continued clonal regression and others that promote clonal expansion. This lends support to therapeutic strategies aiming at complete CLL eradication, particularly in patients with high-risk biological features.

Two findings from our study are of particular interest. Firstly, IgM BCR signaling promotes CLL survival and growth,[33] and hence its reduction to unmeasurable levels in spontaneously regressing cases suggests a quiescent, possibly anergic state that may lead irreversibly to cell death. Interestingly, although measurable signals remained, sIgD expression and IgD signaling were also reduced in the residual circulating tumor cells from

spontaneously regressed CLL patients. While the significance of these changes remains to be determined, a similar reduction of sIgD has been observed in circulating CLL cells of patients receiving ibrutinib therapy that were unable to return to their protective niche within lymph nodes. In these cells, the reduced sIgD levels and IgD signaling capacity were likely consequent to ongoing autophagy, possibly reflecting an ultimate attempt of the CLL cells to survive.[50] Secondly, CLL cells are thought to be resistant to Fas-mediated apoptosis,[51,52] in part through overexpression of TOSO which inhibits procaspase 8 activation.[53] The significance of high FasR expression in spontaneously regressing CLL is therefore intriguing. We did not find a difference in TOSO expression between spontaneously regressed and indolent M-CLL from our RNA-seq analysis. However, there remains a possibility that spontaneously regressing CLL could have an intrinsic sensitivity to Fas-mediated apoptosis. Alternatively, absent IgM signaling may potentially contribute to an increased sensitivity, and future studies will need to clarify this.

Our study identified multiple determinants of spontaneous clonal regression, some of which are intrinsic to the CLL clone (e.g. genetics), whereas others appear to be acquired (e.g. BCR unresponsiveness). Important questions remain as to how the acquired features are developed, and why some CLL tumors, but not others, develop these. One possibility relates to the divergent outcome of different BCR-antigen interactions, some of which may produce sustained responses, while others may elicit responses that decline over time with repeated interactions.[54] Another possibility may reflect potential differences in immunogenicity, with more immunogenic tumors eliciting stronger T-cell responses that result in clonal attrition. Although our preliminary investigation did not reveal significant differences in the expression of T-cell co-stimulatory molecules such as CD80, CD86 and CD40 between regressing and non-regressing CLL, further studies are required to explore these hypotheses.

Finally, complete and partial spontaneous regression cases displayed similar biological features. This suggests that partial regression cases may represent an earlier stage of regression, and with time, complete remission could become apparent. In conclusion, our

study into the biological basis of spontaneous regression provides a unique perspective on the processes underpinning CLL pathogenesis, with treatment implications.

Acknowledgements

This work was supported by a Leukaemia & Lymphoma Research (Bloodwise) clinical research training fellowship (13059) (M.K.) and a National Institute for Health Research clinical lectureship (M.K.). Bioinformatics analysis was supported by a Bloodwise programme grant (11045) (T.S.). Telomere length analysis was supported by a Cancer Research UK grant (C17199/A18246) (D.M.B). The authors thank all patients at Queen Elizabeth Hospital Birmingham, Birmingham Heartlands Hospital, St. James's University Hospital and Pinderfields General Hospital who participated in this study. The authors also acknowledge Dr Ingo Ringshausen and Prof. Mark Drayson for discussion and critical review of this work, and Mr Jamie Beaumont and Mr Jusveen Rai for technical support.

Authorship Contribution

M.K. conceived, designed and directed this study. T.S., P.H., P.M. and A.C.R. provided mentorship and guidance to M.K., and co-supervised this work. M.K., A.C.R., C.O., R.E.J., S.D., A.A., P.E., E.S., S.D., J.M., N.G. and M.H. performed experimental work and/or data analysis. A.G., G.P., A.S-O., R.H. and J-B.C. performed bioinformatics analysis. A.C.R., N.D., A.D.B., M.T., F.F., and D.M.B. provided scientific expertise. M.K., H.P., T.M., P.M., S.P., G.P., P.M. and P.H. recruited patients and/or provided patient care. M.K. wrote the manuscript. All authors revised and approved the final version of the manuscript.

Conflict of Interests Disclosures

The authors do not have competing financial interests.

References

1. Diede SJ. Spontaneous regression of metastatic cancer: learning from neuroblastoma. *Nat Rev Cancer*. 2014;14(2):71-72.

2. Paul R, Remes K, Lakkala T, Pelliniemi TT. Spontaneous remission in acute myeloid leukaemia. *Br J Haematol*. 1994;86(1):210-212.

3. Daccache A, Kizhakekuttu T, Siebert J, Veeder M. Hematologic and cytogenetic spontaneous remission in acute monocytic leukemia (FAB M5b) with trisomy 8. *J Clin Oncol*. 2007;25(3):344-346.

4. Musashi M, Abe S, Yamada T, et al. Spontaneous remission in a patient with chronic myelogenous leukemia. *N Engl J Med*. 1997;336(5):337-339.

5. Puig N, Trudel S, Keats JJ, et al. Spontaneous remission in a patient with t(4;14) translocation multiple myeloma. *J Clin Oncol*. 2009;27(33):e194-197.

6. Kaufmann Y, Many A, Rechavi G, et al. Brief report: lymphoma with recurrent cycles of spontaneous remission and relapse--possible role of apoptosis. *N Engl J Med*. 1995;332(8):507-510.

7. Thomas R, Ribeiro I, Shepherd P, et al. Spontaneous clinical regression in chronic lymphocytic leukaemia. *Br J Haematol*. 2002;116(2):341-345.

8. Del Giudice I, Chiaretti S, Tavolaro S, et al. Spontaneous regression of chronic lymphocytic leukemia: clinical and biologic features of 9 cases. *Blood*. 2009;114(3):638-646.

9. Herishanu Y, Solar I, Ben-Ezra J, et al. Complete spontaneous regression of chronic lymphocytic leukemia. *J Clin Oncol*. 2012;30(26):e254-256.

10. Nakhla PS, Butera JN, Treaba DO, Castillo JJ, Quesenberry PJ. Spontaneous regression of chronic lymphocytic leukemia to a monoclonal B-lymphocytosis or to a normal phenotype. *Leuk Lymphoma*. 2013;54(8):1647-1651.

11. Schmidt HH, Sill H, Eibl M, et al. Hodgkin's disease developing after spontaneous remission of chronic lymphocytic leukemia. *Ann Hematol*. 1995;71(5):247-252.

12. Kipps TJ, Stevenson FK, Wu CJ, et al. Chronic lymphocytic leukaemia. *Nat Rev Dis Primers*. 2017;3:16096.

13. Messmer BT, Messmer D, Allen SL, et al. In vivo measurements document the dynamic cellular kinetics of chronic lymphocytic leukemia B cells. *J Clin Invest*. 2005;115(3):755-764.

14. Stevenson FK, Krysov S, Davies AJ, Steele AJ, Packham G. B-cell receptor signaling in chronic lymphocytic leukemia. *Blood*. 2011;118(16):4313-4320.

15. Herishanu Y, Perez-Galan P, Liu D, et al. The lymph node microenvironment promotes B-cell receptor signaling, NF-kappaB activation, and tumor proliferation in chronic lymphocytic leukemia. *Blood*. 2011;117(2):563-574.

16. Landau DA, Carter SL, Stojanov P, et al. Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell*. 2013;152(4):714-726.

17. Landau DA, Tausch E, Taylor-Weiner AN, et al. Mutations driving CLL and their evolution in progression and relapse. *Nature*. 2015;526(7574):525-530.

18. Puente XS, Bea S, Valdes-Mas R, et al. Non-coding recurrent mutations in chronic lymphocytic leukaemia. *Nature*. 2015;526(7574):519-524.

19. Forconi F, Moss P. Perturbation of the normal immune system in patients with CLL. *Blood*. 2015;126(5):573-581.

20. Ramsay AG, Johnson AJ, Lee AM, et al. Chronic lymphocytic leukemia T cells show impaired immunological synapse formation that can be reversed with an immunomodulating drug. *J Clin Invest*. 2008;118(7):2427-2437.

21. Riches JC, Davies JK, McClanahan F, et al. T cells from CLL patients exhibit features of T-cell exhaustion but retain capacity for cytokine production. *Blood*. 2013;121(9):1612-1621.

22. Nunes C, Wong R, Mason M, Fegan C, Man S, Pepper C. Expansion of a CD8(+)PD-1(+) replicative senescence phenotype in early stage CLL patients is associated with inverted CD4:CD8 ratios and disease progression. *Clin Cancer Res*. 2012;18(3):678-687.

23. Burger JA. Nurture versus nature: the microenvironment in chronic lymphocytic leukemia. *Hematology Am Soc Hematol Educ Program*. 2011;2011:96-103.

24. Lin TT, Letsolo BT, Jones RE, et al. Telomere dysfunction and fusion during the progression of chronic lymphocytic leukemia: evidence for a telomere crisis. *Blood*. 2010;116(11):1899-1907.

25. Rosenquist R, Ghia P, Hadzidimitriou A, et al. Immunoglobulin gene sequence analysis in chronic lymphocytic leukemia: updated ERIC recommendations. *Leukemia*. 2017;31(7):1477-1481.

26. Skowronska A, Austen B, Powell JE, et al. ATM germline heterozygosity does not play a role in chronic lymphocytic leukemia initiation but influences rapid disease progression through loss of the remaining ATM allele. *Haematologica*. 2012;97(1):142-146.

27. Rawstron AC, Fazi C, Agathangelidis A, et al. A complementary role of multiparameter flow cytometry and high-throughput sequencing for minimal residual disease detection in chronic lymphocytic leukemia: an European Research Initiative on CLL study. *Leukemia*. 2016;30(4):929-936.

28. Murphy EJ, Neuberg DS, Rassenti LZ, et al. Leukemia-cell proliferation and disease progression in patients with early stage chronic lymphocytic leukemia. *Leukemia*. 2017;31(6):1348-1354.

29. Vlad A, Deglesne PA, Letestu R, et al. Down-regulation of CXCR4 and CD62L in chronic lymphocytic leukemia cells is triggered by B-cell receptor ligation and associated with progressive disease. *Cancer Res*. 2009;69(16):6387-6395.

30. Calissano C, Damle RN, Marsilio S, et al. Intraclonal complexity in chronic lymphocytic leukemia: fractions enriched in recently born/divided and older/quiescent cells. *Mol Med*. 2011;17(11-12):1374-1382.

31. Calissano C, Damle RN, Hayes G, et al. In vivo intraclonal and interclonal kinetic heterogeneity in B-cell chronic lymphocytic leukemia. *Blood*. 2009;114(23):4832-4842.

32. Damle RN, Batliwalla FM, Ghiotto F, et al. Telomere length and telomerase activity delineate distinctive replicative features of the B-CLL subgroups defined by immunoglobulin V gene mutations. *Blood*. 2004;103(2):375-382.

33. Guarini A, Chiaretti S, Tavolaro S, et al. BCR ligation induced by IgM stimulation results in gene expression and functional changes only in IgV H unmutated chronic lymphocytic leukemia (CLL) cells. *Blood*. 2008;112(3):782-792.

34. Lanham S, Hamblin T, Oscier D, Ibbotson R, Stevenson F, Packham G. Differential signaling via surface IgM is associated with VH gene mutational status and CD38 expression in chronic lymphocytic leukemia. *Blood*. 2003;101(3):1087-1093.

35. Mockridge CI, Potter KN, Wheatley I, Neville LA, Packham G, Stevenson FK. Reversible anergy of sIgM-mediated signaling in the two subsets of CLL defined by VH-gene mutational status. *Blood*. 2007;109(10):4424-4431.

36. D'Avola A, Drennan S, Tracy I, et al. Surface IgM expression and function are associated with clinical behavior, genetic abnormalities, and DNA methylation in CLL. *Blood*. 2016;128(6):816-826.

37. Coelho V, Krysov S, Steele A, et al. Identification in CLL of circulating intraclonal subgroups with varying B-cell receptor expression and function. *Blood*. 2013;122(15):2664-2672.

38. Muzio M, Apollonio B, Scielzo C, et al. Constitutive activation of distinct BCR-signaling pathways in a subset of CLL patients: a molecular signature of anergy. *Blood*. 2008;112(1):188-195.

39. Perbellini O, Falisi E, Giaretta I, et al. Clinical significance of LAIR1 (CD305) as assessed by flow cytometry in a prospective series of patients with chronic lymphocytic leukemia. *Haematologica*. 2014;99(5):881-887.

40. Till KJ, Spiller DG, Harris RJ, Chen H, Zuzel M, Cawley JC. CLL, but not normal, B cells are dependent on autocrine VEGF and alpha4beta1 integrin for chemokine-induced motility on and through endothelium. *Blood*. 2005;105(12):4813-4819.

41. Brachtl G, Pinon Hofbauer J, Greil R, Hartmann TN. The pathogenic relevance of the prognostic markers CD38 and CD49d in chronic lymphocytic leukemia. *Ann Hematol*. 2014;93(3):361-374.

42. Gu B, Dao LP, Wiley J. Impaired transendothelial migration of B-CLL lymphocytes: a defect linked to low L-selectin expression. *Leuk Lymphoma*. 2001;42(1-2):5-12.

43. Yu J, Chen L, Cui B, et al. Wnt5a induces ROR1/ROR2 heterooligomerization to enhance leukemia chemotaxis and proliferation. *J Clin Invest*. 2016;126(2):585-598.

44. Francia di Celle P, Mariani S, Riera L, Stacchini A, Reato G, Foa R. Interleukin-8 induces the accumulation of B-cell chronic lymphocytic leukemia cells by prolonging survival in an autocrine fashion. *Blood*. 1996;87(10):4382-4389.

45. Binsky I, Haran M, Starlets D, et al. IL-8 secreted in a macrophage migration-inhibitory factor- and CD74-dependent manner regulates B cell chronic lymphocytic leukemia survival. *Proc Natl Acad Sci U S A*. 2007;104(33):13408-13413.

46. Burgess M, Cheung C, Chambers L, et al. CCL2 and CXCL2 enhance survival of primary chronic lymphocytic leukemia cells in vitro. *Leuk Lymphoma*. 2012;53(10):1988-1998.

47. Barrio S, Shanafelt TD, Ojha J, et al. Genomic characterization of high-count MBL cases indicates that early detection of driver mutations and subclonal expansion are predictors of adverse clinical outcome. *Leukemia*. 2017;31(1):170-176.

48. Guo B, Zhang L, Chiorazzi N, Rothstein TL. IL-4 rescues surface IgM expression in chronic lymphocytic leukemia. *Blood*. 2016;128(4):553-562.

49. Aguilar-Hernandez MM, Blunt MD, Dobson R, et al. IL-4 enhances expression and function of surface IgM in CLL cells. *Blood*. 2016;127(24):3015-3025.

50. Drennan S, Chiodin G, D'Avola A, et al. Ibrutinib Therapy Releases Leukemic Surface IgM from Antigen Drive in Chronic Lymphocytic Leukemia Patients. *Clin Cancer Res*. 2019;25(8):2503-2512.

51. Panayiotidis P, Ganeshaguru K, Foroni L, Hoffbrand AV. Expression and function of the FAS antigen in B chronic lymphocytic leukemia and hairy cell leukemia. *Leukemia*. 1995;9(7):1227-1232.

52. Greaney P, Nahimana A, Lagopoulos L, et al. A Fas agonist induces high levels of apoptosis in haematological malignancies. *Leuk Res*. 2006;30(4):415-426.

53. Proto-Siqueira R, Panepucci RA, Careta FP, et al. SAGE analysis demonstrates increased expression of TOSO contributing to Fas-mediated resistance in CLL. *Blood*. 2008;112(2):394-397.

54. Packham G, Krysov S, Allen A, et al. The outcome of B-cell receptor signaling in chronic lymphocytic leukemia: proliferation or anergy. *Haematologica*. 2014;99(7):1138-1148.

Table 1

	Diagnosis		Sex	ALC (x10^9/L)		Medical history and co-morbidities	Second malignancy		Non-CLL related drug history
	Year	Age		Peak	Current		Site	Diagnosis	
Complete spontaneous regression									
CLL01	2005	58	M	9.0	3.2	Osteoarthritis, hypertension, hypercholesterolemia	Prostate	2014	Atorvastatin, amlodipine (discontinued)
CLL02	2001	52	M	29.5	3.3	Hypertension, diabetes mellitus	-	-	Metformin, ramipril, bendroflumethiazide, doxazosin, verapamil
CLL03	1996	66	M	32.7	1.5	Ischemic heart disease, peripheral vascular disease, diabetes mellitus, bronchiectasis (since 2010), hypertension	-	-	Alfuzosin, aspirin, bendroflumethiazide, doxazosin, felodipine, lisinopril, simvastatin
CLL04	1994	58	M	10.5	1.7	Hypertension, benign prostatic hypertrophy, previous neurofibroma, osteoarthritis	-	-	Finasteride, amlodipine, ibersartan
CLL05	2000	56	M	25.7	2.6	Previous rectal polyp (non-malignant), diabetes mellitus	-	-	Metformin
CLL06	1997	61	M	27.4	3.5	Diabetes mellitus, hypertension	-	-	Metformin, furosemide
CLL07	2002	83	M	23.2	3.6	Hypothyroidism, diabetes mellitus, hypertension, benign prostatic hypertrophy	Skin (SCC)	2010	Aspirin, linagliptin, tamsulosin, levothyroxine
CLL08	2002	72	F	6.7	2.4	Atrial fibrillation, diverticulosis, hypertension, hyperthyroidism, cholecystitis (in 2014)	-	-	Bisoprolol
CLL09	1990	46	M	20.2	3.1	Hypertension, hypercholesterolemia	-	-	Ramipril, simvastatin
CLL10	2008	70	M	13.8	3.2	Hypertension, abdominal aortic aneurysm, previous TIA, hypercholesterolemia	-	-	Simvastatin, aspirin, lansoprazole, indapamide
CLL11	2007	75	M	14.5	1.2	Atrial fibrillation, gout	Skin (SCC)	2013	Warfarin, bisoprolol, lansoprazole, febuxostat, donepezil
Partial spontaneous regression									
CLL12	2000	66	M	40.6	5.6	Previous iron deficiency anemia	Skin (BCC)	2010	Omeprazole, temazepam, aspirin
CLL13	2013	73	M	161.0	11.3	Abdominal aortic aneurysm, COPD	-	-	Tiotropium and salbutamol inhalers
CLL14	2002	73	F	44.5	8.5	Ischemic heart disease, osteoarthritis, hypertension, hypercholesterolemia	-	-	Aspirin, ramipril, furosemide, ropinirole, levetiracetam
CLL15	2004	61	M	25.9	5.5	Diabetes mellitus, hypertension, hypercholesterolemia			Simvastatin, doxazosin
CLL16	2007	36	F	19.6	7.0	Eczema, mild depression	-	-	Nil
CLL17	2009	81	M	42.4	16.9	Lewy body dementia, previous ischemic heart disease, depression	-	-	Bumetanide, citalopram, donepezil
Spontaneous regression occurring prior to a second malignancy									
CLL18	2008	60	M	50.2	3.0	Diabetes mellitus, hypertension	Prostate	2010	Aspirin, enalapril, metformin, pioglitazone, simvastatin, tamsulosin
CLL19	1987	55	M	78.0	7.7	Diverticulosis, COPD, gout	Lung	2010	Carbocisteine, tamsulosin, tiotropium/fluticasone/ salmeterol inhalers
Relapsed spontaneous regression									
CLL20	1993	38	M	52.4	20.0	Hypercholesterolemia, asthma, previous hip replacement	-	-	Simvastatin, salbutamol/formoterol inhalers

Clinical features of subjects with spontaneous CLL regression. ALC, absolute lymphocyte count; TIA, transient ischemic attack; COPD, chronic obstructive pulmonary disease; SCC, squamous cell carcinoma; BCC, basal cell carcinoma.

Table 2

	B-cell count x10^9/L	CLL count x10^9/L	% CLL cells of B-cells	% CLL cells of lympho-cytes	*IGH* gene usage			*IGHV* % homology to germline	T-cell count x10^9/L	CD4+ T-cell count x10^9/L	CD8+ T-cell count x10^9/L	NK-cell count x10^9/L
					VH	*DH*	*JH*					
CLL01	1.0	0.9	92.0	28.8	3-30	4-17	4	95.5	1.9	1.2	0.7	0.3
CLL02	2.1	2.1	99.0	63.6	3-53	2-21	4	93.7	1.0	0.6	0.4	0.2
CLL03	0.4	0.4	91.6	23.8	1-69	5-12	3	91.7	0.8	0.4	0.4	0.3
CLL04	0.5	0.5	91.5	28.0	3-48	5-18	1	95.5	1.0	0.7	0.3	0.2
CLL05	1.4	1.0	71.6	38.5	2-5	2-15	4	96.3	1.1	0.5	0.4	0.1
CLL06	1.7	1.7	98.8	48.0	4-34	6-13	4	87.9	1.7	0.6	1.0	0.1
CLL07	2.2	2.2	98.0	59.9	3-23	6-19	3	93.1	1.2	0.7	0.5	0.2
CLL08	-	-	86.1	-	3-15	1-1	3	89.1	-	-	-	-
CLL09	2.0	1.8	92.2	58.0	4-39	2-8	5	91.8	0.7	0.5	0.2	0.4
CLL10	1.3	1.2	92.5	37.6	3-13	3-16	6	92.6	1.8	1.0	0.7	0.1
CLL11	0.7	0.7	98.3	57.3	4-4	3-22	4	94.8	0.4	0.3	0.2	0.1
CLL12	1.6	1.2	75.0	21.4	3-30	3-10	4	91.7	3.9	1.0	2.7	0.1
CLL13	5.3	5.3	99.8	46.8	3-74	2-21	5	88.9	4.4	3.6	0.6	1.6
CLL14	6.6	6.5	98.2	76.2	3-23	6-6	2	89.9	1.6	0.8	0.6	0.3
CLL15	3.4	3.4	99.2	61.3	3-30	1-1	3	97.6	1.8	0.9	0.9	0.3
CLL16	6.1	5.9	96.4	84.0	3-23	4-17	5	92.7	0.6	0.3	0.3	0.3
CLL18	1.9	1.8	92.4	58.5	3-23	6-19	5	93.1	0.9	0.4	0.4	0.3
CLL19	6.7	6.2	92.0	80.0	4-34	5-18	6	95.1	0.7	0.4	0.3	0.3
CLL20	-	-	-	-	4-34	6-19	3	94.7	-	-	-	-

Lymphocyte count and characteristics of the *IGH* gene in subjects with spontaneous CLL regression. B-cell, CD4+ T-cell, CD8+ T-cell and NK-cell count were derived from the ALC taking into consideration the proportion of lymphocytes that were CD19+, CD3+CD4+, CD3+CD8+ and CD16+ CD3-neg CD4-neg respectively. The CLL count was derived from the B-cell count taking into account of the proportion of CLL cells in the B-lymphocyte population, determined using the methodology illustrated in Figure 2A. *IGH* sequencing was performed on DNA from sorted CLL cells using *IGHV* leader and *IGHJ* consensus primers. Gene usage and % homology to the germline *IGHV* sequence was determined using the international immunogenetics information (IGMT) platform.

Table 3

Gene name	logFC	*P*-value	Function
CXCL8	3.20	0.001	Chemokine
CXCL7	2.72	0.001	Chemokine
CXCL2	2.66	0.002	Chemokine
CXCL4	2.52	0.002	Chemokine
CCR1	2.03	<0.001	Chemokine receptor
CXCL4L1	1.74	0.001	Chemokine
CXCL5	1.34	0.003	Chemokine

Differentially expressed genes encoding chemokines and chemokine receptors in spontaneously regressed CLL compared with indolent M-CLL. Transcriptome sequencing data from sorted CD19+CD5+ CLL cells was analyzed comparing spontaneously regressed CLL (n=15) against indolent M-CLL (n=17). Positive logFC values identify genes that are upregulated in spontaneously regressed CLL.

Table 4

	CNV
CLL01	8q24.13 del (x1; *TRIB1*; *NSMCE2*); 13q14.2- q14.3 del (x1; miR15a/16-1, *DLEU7; RNASEH2B)*
CLL02	**10q26.3 gain** (x3); 13q14.2- q14.3 del (x1; miR15a/16-1, *DLEU7; RNASEH2B*)
CLL03	13q14.2- q14.3 del (x1; miR15a/16-1, *DLEU7*)
CLL04	2q13 del (x1); **3p24.2-p24.3 gain** (x3; *NKIRAS1*)
CLL05	13q14.2- q14.3 del (x0; miR15a/16-1, *DLEU7*)
CLL06	**3q23 gain** (x3); 13q14.2-q14.3 del (x0; miR15a/16-1, *DLEU7; RNASEH2B*)
CLL07	3q11.1-q11.2 gain (x3) ; **4q28.3 gain** (x3); **11p13-p14.1 gain** (x3; *WT1*)
CLL08	No detectable CNV
CLL09	**7q11.23 del** (x1)
CLL10	7q31.33-q32.1 gain (x3; *POT1*); 13q14.2- q14.3 del (x1; miR15a/16-1, *DLEU7; RNASEH2B*)
CLL11	13q14.2-q14.3 del (x1; miR15a/16-1, *DLEU7*)
CLL12	2p16.1 del (x1)
CLL13	8p del (x1; *TNFRSF10A*); 13q14.2-q14.3 del (x1; miR15a/16-1, *DLEU7*); 17p del (x1; *TP53*)
CLL15	**12q21.2-q31 gain** (x3; *PAWR*); 13q14.2-q14.3 del (x1; miR15a/16-1, *DLEU7; RNASEH2B*); **16q11.2-q12.1 gain** (x3)
CLL16	13q14.2-q14.3 del (x1; miR15a/16-1, *RNASEH2B*; *RB1*)
CLL18	6p22.3 gain (x3)
CLL19	**7p12.1 gain** (x3)
CLL20	13q14.2-q14.3 del (x1; miR15a/16-1, *DLEU7; RNASEH2B*)

Copy number variations (CNV) in 18 spontaneously regressed CLL tumors. Shown in brackets following each CNV are the copy number within the cytoband, followed by any genes within the cytoband that could potentially be functionally significant. Highlighted in bold are CNVs that have not been previously reported in major CLL genomic studies. Data derived from the sorted CD19+CD5+ CLL population.

Table 5

Gene name	logFC	*P*-value	Function
SETMAR	0.81	0.002	Component of the telomerase complex essential for telomere maintenance
FANCL	0.71	0.002	Mediates monoubiquitination of FANCD2, an important step in DNA damage response
MLH3	0.68	<0.001	Important component of DNA mismatch repair
RAD50	0.67	<0.001	Essential mediator of double-strand break repair, DNA recombination and telomere maintenance
ALKBH8	0.58	0.001	tRNA nucleoside involved in DNA damage response
POLA1	0.58	0.001	Regulates the initiation of DNA replication
PRIM2	0.57	0.001	Synthesizes the small RNA primers used to generate Okazaki fragments in the lagging strand during DNA replication
ATM	0.56	0.002	Essential mediator in double-strand break repair, DNA recombination and apoptosis in response to genotoxic stress
MMS22L	0.52	0.003	Involved in the repair of stalled or collapsed replication forks by recombination
ATRX	0.48	0.002	Involved in gene regulation and chromatin modeling during DNA replication
POT1	0.44	0.003	Component of the telomerase complex essential for telomere maintenance
APEX1	-0.68	<0.001	Involved in cellular response to oxidative stress

Differentially expressed genes involved in DNA repair and genome stability in spontaneously regressed CLL compared with indolent M-CLL. Transcriptome sequencing data derived from sorted CD19+CD5+ CLL cells was analyzed comparing spontaneously regressed CLL (n=15) against indolent M-CLL (n=17). Positive logFC values identify genes that are upregulated in spontaneously regressed CLL, whereas negative logFC values identify genes that are downregulated.

Figure legends

Figure 1. Clinical features of 20 spontaneous CLL regression cases. Spontaneous CLL regression cases were categorized into 4 groups: complete spontaneous regression (red curves), partial spontaneous regression (blue curves), spontaneous regression occurring prior to a second malignancy (green curves) and relapsed spontaneous regression (purple curves). The absolute lymphocyte count (ALC) for each patient is plotted against the time elapsed from the date of diagnosis. Clinical features are annotated, and the time of peripheral blood sampling is displayed at the bottom each chart. T0 represents the diagnostic timepoint, whereas T1 and T2 represent the regression timepoints, except in CLL20 where T2 represents the relapse timepoint. LN, lymphadenopathy (the measurement displayed corresponds to the size of the largest palpable node); Tx, treatment

Figure 2. Spontaneously regressed CLL tumors are composed of a non-proliferating CLL clone that has previously undergone substantial cell division. (A) A CD19 gate was applied to the PBMC cells to select for CD19+ B-lymphocytes (not shown). The CLL population (shown in red) could be distinguished from the non-malignant B-lymphocyte population (shown in green) by their immunophenotype, as demonstrated by the example of CLL05. CLL cells exhibit high CD5 and CD43 but low CD20, CD79b and CD81 expression, whereas non-malignant B-lymphocytes exhibit high CD20, CD79b and/or CD81 expression. This methodology was used to determine the proportion of CLL cells within the B-lymphocyte population, the results of which are displayed in Table 2. (B-C) The gated CLL population of PBMCs was analyzed for the expression of Ki-67 and CXCR4. For each comparison, 17 spontaneous regression cases (REG) from the regression timepoint (T1) were compared against 54 indolent (INDOL) and 40 progressive (PROG) cases. Complete and partial spontaneous regression cases are represented by red and blue dots respectively as shown. Statistical significance was determined using one-way ANOVA with Bonferroni post-hoc analysis for cohort comparison and two-tailed paired student t-test for timepoint comparison. Statistical significance is indicated by *$p<0.05$, **$p<0.01$ and ****$p<0.0001$; ns denotes

comparisons that are not statistically significant. (D) Expression of Ki-67 and CXCR4 was compared between two sequential timepoints in individual spontaneous regression cases. T0 and T1 represent the diagnostic and the regression timepoint respectively. Each colored line represents a specific case as indicated. (E) XpYp STELA was performed on sorted CD19+CD5+ CLLs to assess the telomere length of chromosomes Xp and Yp of CLL cells. Left panel: The telomere length distribution of the CD19+CD5+ sorted CLL Xp and Yp chromosomes in each case is shown. Right panel: The mean CLL XpYp telomere length was compared between spontaneously regressed and indolent cases. Complete and partial spontaneous regression cases are represented by red and blue dots respectively as shown. Statistical significance was determined using Student's T test; n.s. denotes comparisons that are not statistically significant.

Figure 3. Spontaneously regressed CLL tumors are characterized by weak BCR signaling response and sIgM expression. FACS analysis of the gated CLL fraction of PBMCs. BCR signaling responses to IgM and IgD stimulation was assessed by phosphoflow using fresh blood samples. Cells were stimulated with anti-human IgM and/or IgD $F(ab')_2$ antibodies prior to acquisition on a flow cytometer. Combined IgM and IgD stimulation reflects BCR stimulation *in* vivo, whereas separate IgM and IgD stimulation allows dissection of the relative contribution of IgM and IgD responses to the overall BCR signaling response in each comparator group. (A) Left panel: The CD19+ CD5+ CLL population (shown in red) was gated and analyzed for the phosphorylation of Syk, Erk and Akt, with the CD19− CD5+ T-cell population (shown in blue) being used as the internal negative control. The positive vs negative gate for p-Syk, p-Erk and p-Akt was set such that 99% of unstimulated cells would fall within the negative gate. Right panel: Example histograms showing results of a typical indolent M-CLL case and a spontaneous regression case with combined IgM/IgD stimulation. Phosphoprotein response to (B) IgM, (C) IgD or (D) combined IgM and IgD BCR stimulation, as well as CLL cell surface (E) IgM (sIgM) and (F) IgD (sIgD) expression. Spontaneous regression cases (REG; n=14) from the regression timepoint (T1) were compared against

indolent (INDOL) M-CLL (n=35) and UM-CLL (n=4) cases. Complete and partial spontaneous regression cases are represented by red and blue dots respectively as shown. (G) Expression of sIgM and sIgD was compared between two sequential timepoints in individual spontaneous regression cases. T0 and T1 represent the diagnostic and the regression timepoint respectively. Each colored line represents a specific case as indicated. Statistical significance was determined using one-way ANOVA with Bonferroni post-hoc analysis for cohort comparison and two-tailed paired student t-test for timepoint comparison. Statistical significance is indicated by *p<0.05, **p<0.01, ***p<0.001 and ****p<0.0001; ns denotes comparisons that are not statistically significant.

Figure 4. Spontaneously regressed CLL tumors express weak surface CD49d but retain high levels of Bcl-2. FACS analysis of the gated CLL population of PBMCs. The gated CLL population was analyzed for (A) basal Erk and Akt phosphorylation and (B) cell surface LAIR1 expression. Spontaneous regression cases (REG; n=14) from the regression timepoint (T1) were compared against indolent (INDOL) M-CLL (n=35) and UM-CLL (n=4) cases. Basal Erk and Akt phosphorylation in CLL cells was normalized to B-cells from 3 age-matched healthy donors, and expressed as fold change compared to these controls. CLL expression of (C) ROR1, (D) CD38, (E) CD62L and (F) CD49d on spontaneous regression cases (REG, n=17) from the regression timepoint (T1) compared to indolent (INDOL, n=54) and progressive (PROG, n=40) cases. (G) Expression of CD49d and CD38 was compared between sequential diagnostic (T0) and regression (T1) timepoints in individual spontaneous regression cases. The gated CLL population was analyzed for the expression of (H) FasR, (I) Bcl-2 and (J) Mcl-1. For all comparisons except Mcl-1, 17 spontaneous regression cases (REG) from the regression timepoint (T1) were compared against 54 indolent (INDOL) and 40 progressive (PROG) CLL cases. For Mcl-1, 11 spontaneous regression cases from the regression timepoint (T1) were compared against 20 indolent and 29 progressive cases. In (A-F) and (H-J), complete and partial spontaneous regression are represented by red and blue dots respectively. Statistical significance is indicated by *p<0.05, **p<0.01, ***p<0.001 and

****p<0.0001; ns denotes comparisons that are not statistically significant. In (G), each colored line represents a specific case and 5% expression is indicated (- - -).

Figure 5. Spontaneously regressed CLL tumors exhibit a distinct transcriptomic profile. RNA-seq was carried out on sorted CD19+CD5+ CLL cells from subjects with spontaneous CLL regression (n=15, from the regression timepoint), indolent M-CLL (n=16), progressive M-CLL (n=8) and UM-CLL (n=8), or on isolated B-cells from healthy donors (n=3). (A) Gene expression profiles of spontaneously regressed tumors were compared against that of age-matched indolent M-CLL tumors, by one-way ANOVA with Tukey post-hoc analysis. Hierarchical clustering analysis demonstrates differential gene expression between spontaneously regressed and indolent M-CLL. (B) Gene set enrichment analysis showing differential expression of *MYC* and Myc target genes in spontaneously regressed versus indolent M-CLL tumors. (C) Database for Annotation, Visualization, and Integrated Discovery (DAVID) analysis showing enrichment of biological processes in spontaneously regressed CLL relative to indolent M-CLL. Biological processes that are upregulated in spontaneously regressed CLL are indicated in red, whereas those that are downregulated in these tumors are indicated in blue. (D) Multidimensional principal component analysis (PCA) of all samples by partial least squares discrimination (PLS-DA) showing distinct clustering of spontaneous regression cases. (E) Unsupervised hierarchical clustering analysis of all samples showing clustering of spontaneous regression cases overlapping with indolent M-CLL cases.

Figure 6. Genomic landscape of spontaneously regressing CLL tumors. WES data initially analysed using standard bioinformatics criteria, subsequently underwent further selection to minimise artefacts and focus on variants of potential functional significance. The selected variants were predicted to be truncating or frameshift, missense mutations predicted to be deleterious, variant allele frequency ≥20% and could be validated by Sanger or RNA sequencing. (A) The genomic landscape of 18 spontaneously regressed sorted CD19+CD5+CLL tumors is displayed. The CNV data of each spontaneous regression case is

combined with their respective SNV data obtained from WES. Different types of genomic events are represented by different colors, with the color code displayed adjacent to the table. The frequency of each genomic event is represented in the bar chart beneath the table. VAF, variant allelic fraction. (B-D) Fish plots of somatic variants in sorted CD19+CD5+CLL cells from (B) CLL05, CLL06 and CLL19 showing clonal equilibrium during the course of spontaneous regression, (C) CLL20 showing clonal equilibrium followed by clonal evolution during CLL relapse after spontaneous regression, and (D) CLL02 showing clonal fluctuation during regression. T0, T1 and T2 represent the diagnostic, regression and relapse timepoints, respectively.

Figure 7. T-cell phenotypic features of spontaneously regressing tumors, re-activation of BCR signaling associated with relapse following spontaneous CLL regression and biological processes underpinning spontaneous CLL regression. FACS analysis of the gated T-cell population of CLL PBMCs. (A) PD-1, LAG3, CD160, CD244, CTLA4 and Tim3 expression on CD4+ and CD8+ T-cells in spontaneous regression cases (REG; n=16) from the regression timepoint (T1) was compared against indolent cases (INDOL; n=31) and age-matched healthy controls (CTRL; n=6). (B) PD-1 expression in T-cells was compared between the diagnostic (T0) and regression (T1) timepoints in individuals with spontaneous CLL regression. (C) Ki-67 expression in T-cells of spontaneous regression cases (CTRL; n=16) from the regression timepoint (T1) was compared against indolent cases (INDOL; n=31) and age-matched healthy controls (CTRL; n=6) and also between the diagnostic (T0) and regression (T1) timepoints in individuals with spontaneous CLL regression. (D) PBMCs from spontaneously regressed cases at maximal regression (n=7) were incubated at 37°C for 48 hours in the presence and absence of CD40L-expressing fibroblasts and IL-4 (25 ng/mL), after which the CLL population was analyzed for Erk, Akt and Syk phosphorylation in response to combined IgM/IgD BCR stimulation. Expression of (E) Ki-67 and (F) phosphoprotein response to combined BCR stimulation of PBMC with anti-human IgM and IgD $F(ab')_2$ antibodies were compared between the regression (T1) and relapse (T2) timepoints of CLL20. (G) Schematic

diagram illustrating the key biological processes underpinning spontaneous CLL regression (REG) and key differences as compared to non-regressing indolent (INDOL) and progressive (PROG) CLL. Statistical significance was determined using one-way ANOVA with Bonferroni post-hoc analysis between CLL cohorts and paired t-test between timepoints or stimulus. Statistical significance is indicated by *p<0.05, **p<0.01, ***p<0.001 and ****p<0.0001; ns denotes comparisons that are not statistically significant.